

Basic Science Symposium III: Animal Models for Orthopaedic Implant Evaluation

Basic Science Symposia Deputy Editor: Vijay K. Goel, PhD
Animal Models Symposium Editor: Lisa Ferrara, PhD

PANEL MEMBERS:



Matthew J. Allen, Vet MB, PhD

Department of Veterinary Clinical Sciences, College of Veterinary Medicine, The Ohio State University, Columbus, Ohio



Anthony Simon Turner, BVSc, MS, Dipl. ACVS

Department of Clinical Sciences, College of Veterinary Medicine & Biomedical Sciences, Colorado State University, Fort Collins, Colorado



Koichi Sairyo, MD

Department of Orthopedic Surgery School of Medicine, University of Tokushima, Kuramoto, Tokushima, Japan

INTRODUCTION

Lisa Ferrara, PhD^a

Preclinical research is the initial prerequisite to determining the efficacy of an implant for eventual translation into the clinical arena. Once a spinal implant has been validated through benchtop mechanical testing, analytical modeling, and in vitro evaluation, it is often useful to conduct an in vivo study using animal models to assess the interactions of the living tissue environment surrounding the spinal implant. Animal models are often implemented in the assessment of spinal implant behavior to identify potential problems with respect to tissue infiltration and adverse tissue and interface reactions. Biomechanical and biological investigations of the tissue and implant interface, as well as the surrounding supportive tissue structures, can provide early insight into the potential clinical performance once implanted into humans. However, the use of animals for implant evaluation remains controversial due to the varied anatomy, different healing rates, and complicated biomechanical environments. There are numerous challenges that exist with animal models such as determining (1) when an animal study is needed, (2) what the appropriate animal model is, (3) what the appropriate time points and outcome measures are, and (4) what the optimal sample size is. Another challenge for such studies is the need to downsize the spinal implants—a serious issue for some animal models.

A panel of eminent experts has been assembled for this symposium to address a number of these issues and challenges with the use of animal models for understanding implant and tissue interface behavior in a living environment. The experts are leaders in their field, and we are fortunate to have them provide insight into these challenges based on their varied experience in this area. We look forward to the readers' participation and feedback to this symposium.

Key Words: Animal model, arthroplasty, quadrupedal, bipedal, spinal devices, tissue ingrowth, wear. *SAS Journal*. December 2008. 2:195-200. DOI: SASJ-2008-Symposium4

^aOrthoKinetic Technologies, LLC, OrthoKinetic Testing Technologies, LLC

Address correspondence to Lisa Ferrara, PhD, OrthoKinetic Technologies, LLC, 2790 Creekbridge Court, Southport, NC 28461 (e-mail: lisa@orthokintech.com)

Question #1. When assessing arthroplasty, motion-sparing, dynamic spinal stabilization devices, what parameters of evaluation (ie, tissue ingrowth, functional behavior, long-term adverse tissue effects, material behavior, wear generation, neural/tissue impingement, etc.) would you require for an animal study to conduct a safe and efficacious evaluation of a novel arthroplasty device?

Allen: As a minimum, I would want the preclinical evaluation of any dynamic, articulating spinal device to include a long-term assessment of adverse tissue effects, wear generation, and device migration/impingement on neural elements. For cementless fixation, I would also expect to see detailed bone ingrowth data at early (6 to 12 weeks) and late (26 to 52 weeks) time points. The assessment of functional behavior

is, in my opinion, a secondary endpoint in preclinical testing because we are often testing devices that are scaled-down human devices rather than custom devices intended for the particular animal species. Functionality data are more commonly derived from in-vitro mechanical testing in joint simulators.

Turner: I will discuss each of the evaluation parameters you mention.

Tissue ingrowth evaluation is an important parameter whether it be spine implants, or hip and knee arthroplasty research. To measure meaningful endpoints in this area, a laboratory capable of preparing undecalcified histological sections along with the appropriate stains for collagen, osteoid/new bone, bone, cell death/apoptosis, is essential. This also requires a bone pathologist capable of reading and interpreting these sections and presenting data (quantitative and qualitative) by means of histomorphometry.

Functional behavior in large animals such as sheep, goats, and dogs is possible when evaluating various arthroplasties (specifically implants in the hip and knee joints). Quantitative and qualitative data by evaluation of use of the limb is straightforward with veterinary assistance. A grading system that correlates the use of the limb with a numerical score (eg, 1–4) is the norm. More sophisticated laboratories with force plates can deliver even more quantitative data, especially in dogs because they are more easily handled and trained to use a force plate, when compared to the more fractious farm animal. However, functional behavior of large animals in *spine implant research* is very difficult. Many of the implants require extensive muscle dissection for the surgical approach, especially those with dorsal (“posterior”) instrumentation. This in itself is painful and results in varying degrees of postoperative morbidity that cannot be differentiated from the effects of the implant used in the spine surgery. Furthermore, many Institutional Animal Care and Use Committees (IACUC’s) justifiably require preemptive analgesics *and* at least 3 days of postoperative analgesia. This would make functional behavior in animals used for spine implant evaluation even more difficult. Large animals such as sheep, goats, and dogs undergoing ventral (“anterior”) fusions of 1-4 levels of the lumbar spine tolerate the surgery very well and are usually back to “normal” cage/pen activities within 24 hours of surgery. Therefore, functional behavior in these animals following this type of spine surgery is virtually impossible.

Some surgeries of the spine in animals involve close contact with nerve roots or, in the case of laminectomies, the spinal cord itself. These techniques are ones where paralysis/paresis are the most likely of the complications and evaluation of the animal as to its ability to stand, to walk without dragging a limb, or be clinically normal is a real possibility. A pain scoring system is required by many IACUC’s.

Long-term adverse tissue effects as endpoints can be evaluated at the time of necropsy by visualization of the tissues surrounding, or in close proximity to, the implants. Staining of these tissues can be noted macroscopically although simple H&E histology is always indicated. Other tissues in close proximity to the spine that may need to be evaluated histologically are the sub-lumbar lymph nodes (especially if wear debris is anticipated) and paraspinous muscles (*m. longissimus dorsi*, *m. psoas*, *m. multifidus*, etc.). Some safety studies require histological evaluation of representative samples of parenchymatous organs such as the liver, spleen, kidney, heart muscle, lung, etc.). Very rarely, regulatory agencies require representative samples of the central nervous system (brain) or reproductive system (ovary, uterus, vagina, etc.) to be evaluated histologically. These histological sections, when mounted and stained appropriately, are best evaluated by a board-certified veterinary pathologist. There are frequently lesions in domesticated animals that are clinically insignificant or even normal (eg, *Sarcocystis sp.* in the muscles of sheep and goats) that would alarm a pathologist unfamiliar with incidental findings in these tissues in animals.

Material behavior can be determined by bioengineers in the laboratory in vitro, long before the sponsor’s money is spent surgically implanting the implants into large animals in preclinical in vivo experiments. Evaluation of the material and surrounding tissue reaction is determined at necropsy and then histologically. Evaluation of the explanted spines immediately after necropsy is possible in a laboratory equipped with the appropriate materials testing equipment (eg, MTS or Instron systems). Prior to evaluating the explanted spines acquired from animals that are in a study, it is always

recommended that the engineer obtain cadaveric spines from the appropriate animal to perfect the mounting and potting techniques used in the materials testing equipment.

Neural/tissue impingement is difficult to evaluate clinically unless there has been a surgical error in placement of implants.

Sairyo: For devices providing motion preservation, bone-implant interface is the most important factor in obtaining better long-term outcomes. In contrast, for fusion devices such as the pedicle screw rod system, long-term outcomes

depend on bony union, not on the device. This is the biggest difference between motion-preservation devices and fusion devices. Based on these concepts, I think the most important parameter is bone-ingrowth into the surface of the device.

Question #2. What animal models would you pose for each of the parameters addressed in Question 1 and why?

Allen: Whenever possible, I am a proponent of small ruminant models for spine research. In general terms, it should be possible to derive all of the information that is needed from a single animal model. If wear generation is identified as a potential concern in the small ruminant model, additional

studies may be needed in small animal models such as the rat or rabbit in order to more fully characterize the tissue response to different doses of particles. The dose of particles used in those experiments would be guided by data from wear simulator tests.

Turner: The sheep, goat, or larger breeds of dogs are ideal animals for evaluation of spine implants. Rabbits are too small to evaluate most implants that are ultimately to be used in humans. Rabbits are an ideal animal model for the initial evaluation of bone graft substitutes that are used in spine fusion involving decortication of the transverse processes. The advantage of sheep and goats over dogs is their general acceptability as a research animal compared to the dog, especially in Europe and the UK. Pigs have been used, but the weight of certain skeletally mature breeds of domestic pigs makes them impractical to house and handle. Pigs of a reasonable body weight may be skeletally immature. Other

novel breeds such as Yucatan or even “micro-pig” breeds may be acceptable as long as the age of closure of all growth plates is known.

The question arises whether the use of a quadruped can mimic the biomechanical loading of the bipedal spine of humans. Wilke et al.^{1,2} and Smit³ showed that the loading of the quadrupedal spine (sheep) is closer to humans than would be anticipated. There really is no bipedal animal model that ambulates like human beings. Non-human primates use their forearms for support for much of their ambulation.

Sairyo: A larger animal model is better. The size of the spine should be similar to the human spine. Use of a primate model

is optimal, since load transmission of the bony weight is similar to the human.

Question #3. Please express your thoughts on appropriate quadrupedal and/or bipedal animal models to assess the long-term in vivo implant function for the following spinal devices: (A) Total disc arthroplasty, (B) Posterior dynamic stabilization, (C) Facet replacement devices, (D) Interspinous spacers, and (E) Nuclear replacement devices.

Allen:

A. Total disc arthroplasty: Goat, nonhuman primate, and dog would all be appropriate, depending on the size (particularly height) of the device. In our experience, differences in endplate geometry can make stabilization of even a small human device problematic in animals, leading to instability of the device in the immediate postoperative period. Due to the complexity of these models, preclinical studies on total disc replacement function (eg, range of motion, volumetric assessment of implant wear) are probably best conducted in vitro, with animal tests being limited to evaluation of the biological response to the component materials and any associated wear products.

B. Posterior dynamic stabilization: Sheep, pig, or dog are appropriate, depending on whether there is a need for pedicle

screw fixation. If pedicle screws are required, the sheep would be the model of choice since the canine pedicle is extremely narrow, making it hard to insert screws without a high risk of canal intrusion.

C. Facet replacement devices: I am not aware of much literature on the use of these devices in animal models. In order to recreate the size and geometry of the facet joint, devices would have to be custom designed for use in most of the species that are commonly used in spine research.

D. Interspinous spacers: My comments here are similar to those for (C). In-vitro testing is commonly performed on calf spines, but translation into in-vivo models with either calves or miniature cows would be logistically and financially challenging.

E. Nuclear replacement devices: Sheep or nonhuman primate would be appropriate, although our experience has been that it is very hard to ensure that the device is retained within the disc due to problems with (1) high intradiscal pressures, (2) incomplete nucleus removal, and (3) failure to seal the annular incision. Our findings in animals have been

consistent with early human clinical experience with nuclear replacement devices. Significant improvements in annular repair technology will be needed before nuclear replacement can be performed safely in animals or considered a routine clinical procedure in humans.

Turner:

A. Total disc arthroplasty: The endplate dimensions in sheep, goats, and dogs are very different than those of humans, and the only advantage of using animals this size is to gain understanding of the biological behavior of the surrounding tissues to the prosthesis. Specifically, this includes, but is not limited to, the response of host ingrowth at the endplates, the histological evaluation of surrounding paraspinal musculature, spinal cord, and signs of wear debris (foreign body reaction) in surrounding tissues and the sublumbar lymph nodes.

B. Posterior dynamic stabilization: The advantages of sheep, goats, and larger breeds of dog for evaluation of dorsal (“posterior”) dynamic stabilization is that implants used in the human spine can be used most of the time because of the dimensions of the vertebral bodies in these animals. Drastic downsizing of implants would be required for smaller breeds of dog such as the beagle and would be impossible in animals like the rabbit.

C. Facet replacement devices: The size of sheep, goats, and larger breeds of dog would be an advantage although this

seems like a relatively new field of investigation compared to the interbody fusion devices.

D. Interspinous spacers: Interspinous spacers can be readily evaluated in sheep, goats, and large breeds of dog, purely because of size. Typically, 2 interspinous spacers (experimental and predicate or both experimental) can be used with a normal interspace between. The surgery in sheep is well tolerated with low morbidity.

E. Nuclear replacement devices: In my experience, sheep have very small disc “height,” so a certain amount of downsizing is essential for evaluation of any nuclear replacement devices. The disc degeneration model popularized by Australian researchers (see Osti et al.⁴) requires nothing more than a stab incision in the annulus to cause varying degrees of disc degeneration over time. The stab in the annulus required to replace the nucleus pulposus with a device is surely going to cause progressive degeneration of the disc, thus complicating this model.

Sairyo:

A. Total disc arthroplasty: The primate model is better.

B. Posterior dynamic stabilization: The primate model is better.

C. Facet replacement devices: The primate model is better.

D. Interspinous spacers: The primate model is better.

E. Nuclear replacement devices: While the primate model is better, the goat model is acceptable.

Question #4. Please describe the specific challenges regarding the devices listed in Question 3 when using a quadruped animal model to assess functional behavior of an orthopedic implant in the cervical spine and also in the lumbar spine?

Allen: Primary limitations for animal models of total disc replacement would be device size (the height of the device is often greater than available disc/endplate height in the animal model) and range of motion (design of the human implant may not accurately recreate that needed for disc functionality

in the animal model. Differences in load magnitude and frequency may limit the interpretation of wear debris data.

Potential concerns over the remaining devices are discussed in section C (above).

Turner: Many of these challenges have been outlined in my previous responses as they relate to the *lumbar* spine of sheep, goats, and dogs. In the *cervical* spine, most of the studies have involved evaluation of interbody spacers, cages, bone graft substitutes, and osteogenic-stimulating materials

such as growth factors. The vertebral bodies of sheep, goats, and dogs are considerably smaller than those of humans, so smaller sized implants are needed. One thing that is essential with any spinal device that requires a ventral approach to the cervical spine is the need for containment and/or prevention

from expulsion or implant migration. Because of the extreme mobility of the quadruped cervical spine, a spinal device

within the cervical interspace will be rapidly expelled in a ventral direction, without the use of a containment plate.

Sairyo: For the quadruped model of **total disc arthroplasty**, compression loading may not be large enough to simulate the human lumbar spine. However, some information can be acquired with the use of quadruped models.

would be applied to the device. For this kind of device, the quadruped model is not acceptable.

For **posterior dynamic stabilization**, the biped model is better; however, the quadruped is acceptable.

For **interspinous spacers**, the biped model is better; however, for this device, quadruped is acceptable.

In the quadruped model of **facet replacement devices**, stress direction is completely different in the facet joint compared to humans. The artificial facet should be replaced after complete bilateral facetectomy. In quadruped models, high shear force

For modeling **nuclear replacement devices**, it seems that nucleus pulposus (NP) pressure is similar in both quadruped and biped. Thus, the quadruped model is acceptable.

In **implants for the lumbar spine**, loading is more important than it is in the cervical spine implant.

Question #5. What are the drawbacks for utilizing a nonhuman primate model for functional assessment of an arthroplasty device? What are the anatomic and biomechanical similarities and differences between a nonhuman primate model and the human spine?

Allen: Although there are similarities in the anatomy of the human and nonhuman primate (NHP) spine, the use of these species in functional studies is complicated by ethical concerns, biosafety issues health (particularly with respect

to zoonoses such as herpes B in macaques), and financial considerations (NHPs are commonly the most expensive animal model of any that are used in biomedical research).

Turner: The use of non-human primates is important and essential in the regulatory approval of many spine implants. Owing to the expense of such research, primate studies should be done only when other large animal models have been used to their fullest extent. Only a small number of primates may be necessary for evaluation of safety and efficacy of various devices. The smaller primates (eg, *Cynomolgus sp.*) may be of

limited value because of the needed downsizing of implants, and primates the size of baboons (*Papio cynocephalus spp.*) may be required. Housing, veterinary care, social enrichment programs, anesthesia, analgesia, transmission of zoonotic diseases, to name a few, are all hurdles, aside from expense and acceptance to society, that need to be negotiated when dealing with non-human primates.

Sairyo: Since the primate is biped, the biomechanics are very similar. However, the physiological spinal curvature is

different in each, and the biomechanics are not completely alike.

Question #6. What are some of the pitfalls associated with animal studies?

Allen: Potential pitfalls associated with animal studies include:

- Differences in tissue healing and bone remodeling rates can complicate the interpretation of temporal data on implant-bone integration.
- Preclinical data are typically derived from healthy animals and may not be entirely representative of the tissue response in patients with underlying pathology.
- Animal studies usually involve the implantation of human devices that may not be optimally designed to recreate the biomechanics of the animal species.

- Loading patterns may not be the same as those in humans, leading to a more conservative test of device performance and safety.

- Cost constraints may compromise the study design so that group sizes are insufficient to ensure statistical power.

- Ultimately, as in any preclinical study, statistical significance in an animal study may not be predictive of biological significance (ie, true clinical efficacy) in a human patient population.

Turner: Skeletally mature animals are essential because the biological response in young animals can be misleading (unless the target human populations are pediatric or adolescent patients). This means that we use sheep > 3.5 years of age, when all growth plates are radiographically closed and all growth of boney structures has ceased. Animals tend to respond more favorably to fractures (regardless of age), so evaluation of endpoints such as interbody cervical or lumbar

fusion should be done at relatively early time points. For example, many critics of interbody fusion studies in large animals will remark that they “all will eventually fuse.” To measure a biological response to a growth factor, a novel bone graft substitute or fusion device, more meaningful information about the biology may be revealed at early time points (as early as 2 months) rather than waiting for long periods.

Sairyo:

1. For patients, we conduct postsurgical rehabilitation and immobilization with a brace; it is very difficult to do so in the animal model. Animals move when they want to move. That is the biggest difference between patient and animal models.

2. Clinically, surgeons frequently use these implants in degenerated spines that are affected by osteoporosis as well as disc collapse (degeneration) and facet osteoarthritic changes. It is not easy to simulate such degenerated and osteoporotic spines in animal models.

REFERENCES

1. Wilke HJ, Kettler A, Claes LE. Are sheep spines a valid biomechanical model for human spines? *Spine*. 1997;22(20):2365-2374.
2. Wilke HJ, Kettler A, Wenger KH, Claes LE. Anatomy of the sheep spine and its comparison to the human spine. *Anat Rec*. 1997;247(4):542-555.
3. Smit TH. The use of a quadruped as an in vivo model for the study of the spine—biomechanical considerations. *Eur Spine J*. 2002;11:137-144.
4. Osti OL, Vernon-Roberts B, Fraser RD. 1990 Volvo Award in experimental studies. Anulus tears and intervertebral disc degeneration. An experimental study using an animal model. *Spine*. 1990;15(8):762-767.